Available online at www.sciencedirect.com
ScienceDirect
Vision Research 47 (2007) 1282–1291
Vision Research
www.elsevier.com/locate/visres

The lowest spatial frequency channel determines brightness perception

A. Perna [a,b,*], M.C. Morrone [c]

[a] Scuola Normale Superiore, Piazza dei Cavalieri 7, 56100 Pisa, Italy
[b] Istituto di Neuroscienze CNR, via Moruzzi 1, 56127 Pisa, Italy
[c] Università 'Vita-Salute' San Raffaele, via Olgettina 58, 20132 Milan, Italy

Received 27 February 2006; received in revised form 5 January 2007

Abstract

This study investigates the role played by individual spatial scales in determining the apparent brightness of greyscale patterns. We measured the perceived difference in brightness across an edge in the presence of notch filtering and high-pass filtering for two stimulus configurations, one that elicits the perception of transparency and one that appears opaque. For both stimulus configurations, the apparent brightness of the surfaces delimited by the border decreased monotonically with progressive (ideal) high-pass filtering, with a critical cut-off at 1 c/deg. Using two octave ideal notch filtering, the maximum detrimental effect on apparent brightness was observed at about 1 c/deg. Critical frequencies for apparent brightness did not vary with contrast, viewing distance, or surface size, suggesting that apparent brightness is determined by the channel tuned at 1 c/deg. Modelling the data with the local energy model [Morrone, M. C., & Burr, D. C. (1988). Feature detection in human vision: a phase dependent energy model. *Proceedings of the Royal Society (London), B235*, 221–245] at 1 c/deg confirmed the suggestion that this channel mediates apparent brightness for both opaque and transparent borders, with no need for pooling or integration across spatial channels.

Keywords: Brightness; Edges and lines detection; Spatial frequency channels; Filling-in; Feature detection

1. Introduction

The visual system is known to elaborate in parallel many of the attributes of visual images. While parallel analysis guarantees faster processing and stability, it poses the problem of how these independent, sometimes incongruent estimates of scene characteristics can support a unitary explicit perception.

An example of parallel analysis in the visual system is the processing of different spatial frequency bands by independent channels (Campbell & Robson, 1968), providing independent estimates of image attributes at the same retinal location. For near-threshold contrast-levels, scene detection can be simulated by evaluating the probability summation of information across spatial scales (Graham, 1977), implying independent analysis. However, independent analysis may not necessarily hold at supra-threshold contrasts. To predict scene appearance many models assume that the outputs of processing at each scale are summed, with constant weighting independently of the task to be performed (e.g., MIRAGE, Watt & Morgan, 1985; MIDAAS, Kingdom & Moulden, 1992).

At supra-threshold contrast values, contrast appearance does not scale with the size of the stimulus pattern, suggesting that all spatial scales play a similar role in determining the perceptual appearance of contrast and brightness (Georgeson & Sullivan, 1975). However, there are visual scenes whose perception seems to be mediated by a single channel, irrespective to it being the more suitable for the current task or not. Such tasks involve letter identification (Majai, Pelli, Kurshan, & Palomares, 2002; Solomon & Pelli, 1994) and face recognition (Pelli, 1999). Recently, Peromaa and Laurinen (2004) showed that the appearance of a brightness illusion such as Chevreul staircase is medi-

* Corresponding author. Address: UMR 5169, Bat IVR3, 118 route de Narbonne F-31062, Toulouse cedex 09, France. Fax: +39058446249.
E-mail address: perna@cict.fr (A. Perna).

0042-6989/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2007.01.011

ated by the low spatial-frequency components of the pattern, while the visibility of the edges is mainly determined by higher spatial frequencies.

In general, the integration of responses across channels can be described in terms of a Minkowski sum, where the total activity A_{tot} is linked to the activities at the single scales A_i through

$$A_{\text{tot}} = \sqrt[m]{\sum_{i=1}^{s} w_i \cdot |A_i|^m} \tag{1}$$

Depending on the value of m, the equation describes a linear sum of responses across scales, Pythagorean sum, different kinds of probability summation, or winner take all (for very large m). The weights w_i can be equal for all scales, or favour some scales over the others.

An alterative strategy to integration across scales would be to select an optimal channel depending on the task to be performed, for example, on the basis of the strongest response or the highest signal to noise ratio. This type of strategy have been shown to be used by our visual system, at least under some circumstances (Solomon, 2000).

The apparent brightness difference across an edge is known to depend on many factors, such as the luminance profile of the edge (Cornsweet, 1970; Craik, 1966; O'Brien, 1958) and the characteristics of other regions in the visual image, that interfere with brightness in a variety of different ways, including simultaneous contrast (Heinemann, 1955) and assimilation (Shapley & Reid, 1985). In addition, brightness is also influenced by the three-dimensional arrangement of surfaces and image segmentation, both in the presence and in the absence of transparency illusion (Adelson, 1993; Anderson & Winawer, 2005; Singh & Anderson, 2002). In the present experiment, we investigate how the contributions from different spatial frequency channels are integrated to determine the difference in apparent brightness across a luminance-defined edge that give rise to a simultaneous brightness illusion. In order to reach a general finding, we measured both opaque edges and edges that give rise to a transparency illusion.

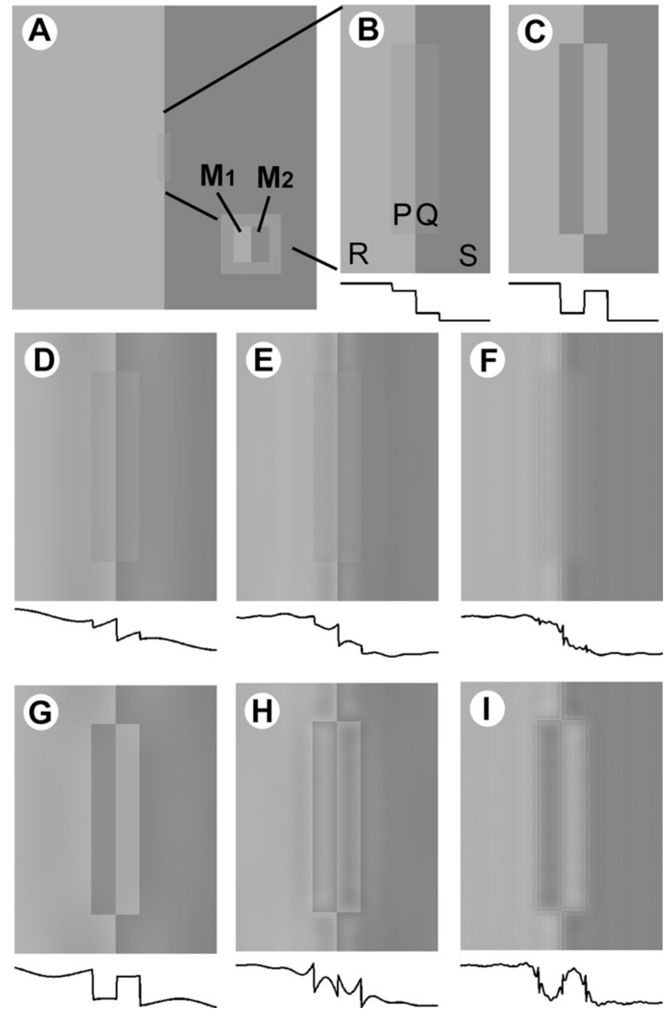

Fig. 1. Configuration of the stimuli used in the experiment. (A) Basic stimulus configuration comprising the test and the match patch. The central region is represented with increased size in (B). Subjects had to compare the contrast of the edge between region P and region Q with the contrast of the edge in the lower right part of the screen (the edge separating $M1$ from $M2$). Luminance of regions R and S were the same for all stimuli, while luminance of regions P and Q was variable (in all the stimuli represented in this figure, P–Q contrast was 17.6%). Opaque configuration stimuli, such as that shown in (C) were obtained inverting contrast polarity of the edge between P and Q, but not of the edge between R and S. Both transparent and opaque stimuli could be spatial filtered in different bands. Examples of the stimuli obtained are shown in (D–I) corresponding to a notch-filter of two octaves positioned at 0.3 (D), 0.9 (E), and 4 (F) c/deg (centre frequency) applied to a "transparent" stimulus and to the same filters applied to an "opaque" stimulus (0.3 c/deg (G), 0.9 c/deg (H), 4 c/deg (I)).

2. General methods

2.1. Apparatus and stimuli

Stimuli were prepared with MATLAB and displayed with a Cambridge Research System VSG 2/3 graphics card on a high-resolution 21 in. BARCO monitor (spatial resolution 656 × 507 pixels, frame refresh rate 126 Hz, mean display luminance 16.6 cd/m^2). The whole stimulus measured 25 × 25 cm, subtending 23.5 × 23.5 deg of visual angle at the viewing distance of 60 cm (with the exception of one experiment where viewing distances of 30 and 120 cm were tested).

The basic stimulus configuration comprising the test and the match patch is illustrated in Fig. 1A. The test stimulus was constructed of two regions of different uniform intensity above and below the mean intensity, marked in the figure by the letters P and Q, and subtended 0.93 w × 3.72 h deg of visual angle. P and Q regions were aligned vertically against a larger pair of background regions of different uniform intensity again above and below the mean intensity R and S (luminance values: S = 11.7 cd/m^2 R = 21.5 cd/m^2, contrast 0.29).

The match stimulus was made up of two regions of different uniform intensity above and below the mean, marked in the figure by the letters $M1$ and $M2$ and subtending 1.38 w × 2.75 h deg of visual angle each. It was located in the right corner of the screen and centred at a distance of 9.5 deg of visual angle from the test patch along the diagonal and was isolated from the other regions by a 0.93 deg wide white paper mask put on the screen to perceptually segregate it from the display (luminance of 8.5 cd/m^2).